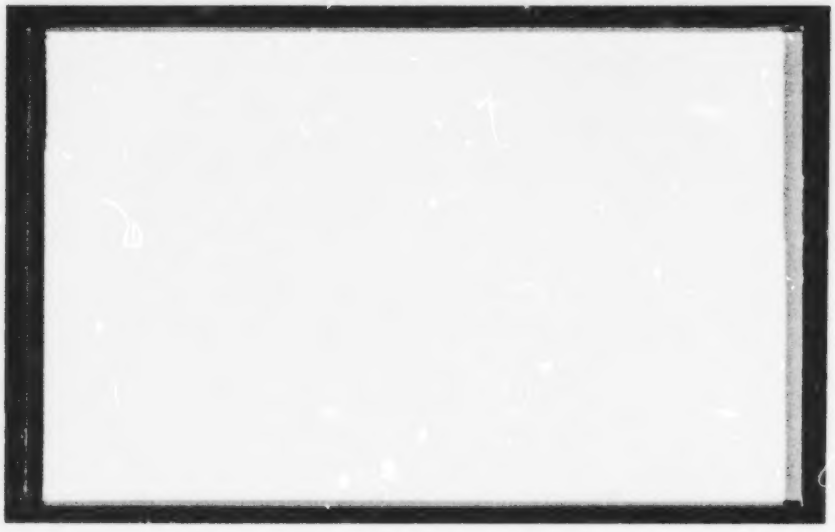
UNIVERSITÉ DE MONTRÉAL



CENTRE DE RECHERCHE SUR LES TRANSPORTS

MÉMOIRE D'ÉTUDIANT

CASE POSTALE 6128, STATION A, MONTRÉAL, QUÉBEC, H3C 3J7

UNE EXPLICATION DU MODE DE PAIEMENT DES DÉPLACEMENTS PAR TRANSPORTS EN COMMUN PAR LES MODÈLES DOGIT ET LOGIT

par

Richard Laliberté

Cette étude a été rendue possible par des subventions du Centre de recherche et de développement de Transports Canada et par le programme F.C.A.C. du Ministère de l'Education du Québec.

Université de Montréal

Centre de recherche sur les transports - Publication # 112

Juillet 1978

Dépôt légal - 4e trimestre 1978

UNIVERSITÉ DE MONTRÉAL

Sommaire

Ce rapport cherche à comparer le dogit au logit pour l'explication de la part relative de marché pour deux altervatives. Nous utilisons des séries chronologiques agrégées sur les modes de paiements des déplacements du transport en commun à Montréal de 1965 à 1971. Par une transformation des variables explicatives du modèles, nous n'incorporons aucune forme fonctionnelle a priori. Nous montrons que le dogit est supérieur au logit; nous montrons par ailleurs que ce résultat est confirmé par la transformation des variables explicatives.

UNIVERSITÉ DE MONTRÉAL

Table des matières

Sommaire iii

1. Introduction 1

2. Spécification du modèle dogit 2

   A. Distribution du modèle 2

   B. Formulation de la forme fonctionnelle 3

3. Propriété du modèle 5

   A. Le logit cas spécial du dogit 5

   B. Libre en général de l'axiôme IIA 5

   C. Possibilité d'un "effet de revenu" 6

   D. Intégration dans une même distribution d'un comportement contraint et libre 7

4. Application du modèle à une série chronologique 9

5. Méthode d'estimation 10

   A. Nouvel arrangement pour les besoins de l'estimation 10

   B. Fonction de vraisemblance en présence d'autocorrélation multiple 12

   C. Technique de maximisation 14

6. Résultats 16

   A. Mesure d'évaluation des résultats 16

   B. Analyse des résultats 18

      B.1. Influence générale des paramètres du dogit 18

      B.2. Comparaison entre les différents modèles 22

7. Conclusion 25

Remerciements 26

Bibliographie 27

Graphique 1 : Valeurs du log de vraisemblance pour Dsc 19

Graphique 2 : Valeurs du log de vraisemblance pour D1 20

Graphique 3 : Valeurs du log de vraisemblance pour D1 21

Tableau 1 : Résultats de la série chronologique 24

I. Introduction

La théorie classique du comportement des consommateurs suppose que les individus réagissent toujours à des variations de prix et de d'autres facteurs. Comme l'a toutefois montré Lancaster (1966), il est possible que ce type d'analyse ne rende pas compte du choix entre altervatives aux caractéristiques multiples. L'une des façons d'expliquer le comportement est alors d'utiliser des modèles probabilistes dont le plus utilisés sont peut-être le probit et le logit (Domencich et Mc Fadden 1975).

Le modèle utilisé ici, le dogit, permet, comme dans le modèle logit, que certains rapports de probabilités de choix entre deux options particulières dépendent exclusivement des attributs de ces dernières mais il permet aussi que d'autres rapports dépendent des attributs de toutes les options considérées.

Les résultats obtenus sont d'un grand intérêt pour ceux qui utilisent les modèles probabilistes dans l'analyse économique du comportement des individus.

2. Formulation du modèle dogit

A. Le modèle en tant que distribution

Considérons la distribution suivante :

$$p_i = \frac{\exp(V_i) + \theta_i \sum_j \exp(V_j)}{(1 + \sum_j \theta_j) \sum_j \exp(V_j)} \quad , \quad i, j = 1, \ldots, N \qquad (1)$$

où p_i : part de la $i^{ième}$ de N alternatives;

V_i : une fonction de K variables indépendantes X_{ik} associée à la $i^{ième}$ alternative;

θ_i : un paramètre positif ou nul associé à la $i^{ième}$ alternative : $\theta_i \geq 0$.

L'équation (1) est défini de telle sorte que

$$0 \leq p_i \leq 1 \qquad \text{où } i = 1, \ldots, N \qquad (2)$$

et

$$\sum_i p_i = \frac{\sum_i \exp(V_i) + \sum_i \theta_i \sum_j \exp(V_j)}{(1 + \sum_j \theta_j) \sum_j \exp(V_j)} = 1 \qquad (3)$$

On peut retrouver la spécification de ce modèle dogit dans le texte de Gaudry et Dagenais (avril 1978).

B. Formulation de la forme fonctionnelle

Gaudry et Wills (juin 1977) ont conclu qu'une mauvaise formulation de la forme fonctionnelle pouvait engendrer de fausses élasticités et de mauvais signes pour des coefficients importants. Nous allons comme eux utiliser une forme fonctionnelle générale.

Dans l'équation (1), la forme fonctionnelle de la fonction de V_i est indéterminée. Nous allons utiliser la transformation suivante :

$$V_i = B_{io} + \sum_k B_k \left[\frac{(X_{ik} + \mu_k)^{\lambda_k} - 1}{\lambda_k} \right] \qquad (4)$$

que nous pouvons réécrire :

$$V_i = B_{io} + \sum_k B_k X_{ik}^{(\lambda_k\ ,\ \mu_k)} \qquad (5)$$

où la transformation $(\lambda_x\ ,\ \mu_x)$ d'une variable x quelconque est définie comme étant :

$$X^{(\lambda_x,\ \mu_x)} = \begin{cases} \dfrac{(X + \mu_x)^{\lambda_x} - 1}{\lambda_x}\ , & \lambda \neq 0\ ,\ X + \mu_x > 0\ , \\ \\ \ell n\ (X + \mu_x)\ ,\ \lambda \to 0\ ,\ X + \mu_x > 0\ . \end{cases} \qquad (6)$$

μ_x est un coefficient de location étudié par Tukey (1957). Box et Cox (1964) ont montré que la transformation est continue pour $\lambda_x = 0$ si nous observons la transformation précédente. Schlesselman (1971) a souligné l'importance de la présence du terme constant en (4) pour préserver l'invariance de la transformation aux unités de mesure.

Selon la forme postulée en (4), si nous posons les vecteurs $\mu = (\mu_1, \ldots, \mu_k)$ et $\theta = (\theta_1, \ldots, \theta_i)$ égaux à 0 nous aurons une forme logit Box-Cox. Si un θ_i n'est pas nul et que tous les vecteurs $\lambda = (\lambda_1, \ldots, \lambda_k)$ sont égaux à 1, nous aurons un dogit additif

$$p_i = \frac{\exp\left(B_{io} + \sum_k B_k X_{ik}\right) + \theta_i \sum_j \exp\left(B_{jo} + \sum_k B_k X_{jk}\right)}{\sum_j \exp\left(B_{jo} + \sum_k B_k X_{jk}\right)\left(1 + \sum_j \theta_j\right)} \tag{7}$$

et si $\lambda = 0$ nous aurons un modèle dogit multiplicatif

$$p_i = \frac{\exp\left[(B_{io}) \prod_k X_{ik}^{B_k}\right] + \theta_i \sum_j \exp\left[(B_{jo}) \prod_k X_{jk}^{B_k}\right]}{\sum_j \exp\left[(B_{jo}) \prod_k X_{jk}^{B_k}\right]\left(1 + \sum_j \theta_j\right)} \tag{8}$$

Nous allons estimer le modèle (1) à l'aide de la transformation V_i décrite à l'équation (4). Etant donné le but de ce rapport, soit la comparaison du dogit face au logit et le grand nombre de modèles implicites définie par (5), nous limiterons à certains cas particuliers des vecteurs μ et λ.

3. Propriétés du modèle

A. Le logit cas spécial du dogit

Si l'on peut montrer, que pour toutes les alternatives, les données se comportent de façon à ce que $\theta_i = \theta_j = 0$ alors dans ce cas précis nous retrouvons le modèle logit suivant :

$$p_i = \frac{\exp (V_i)}{\sum_j \exp (V_j)} , \quad i, j = 1, \ldots, N . \tag{9}$$

B. Libre en général de l'axiôme IIA

Domencich et McFadden (1975) ont démontré que dans le modèle logit, le rapport de probabilités p_i/p_j de choisir l'alternative i sur l'alternative j est indépendant de la présence ou de l'absence d'une autre alternative. Cette conclusion est en accord avec l'axiôme de Luce (1959) sur "l'indépendance des alternatives non pertinentes (IIA; indépendence of irrelevant alternatives)".

Pour leur part, Gaudry et Dagenais (avril 1978) supposent dans le modèle dogit qu'il n'est pas nécessaire que toutes les probabilités relatives soient déterminées simultanément de la même manière. Le dogit peut par ailleurs se comporter comme le logit sans que les θ soient égaux à zéro dans les cas suivants :

$$\frac{O_i}{O_j} = \frac{\exp (V_i)}{\exp (V_j)} , \tag{10}$$

ou

$$\exp (V_j) = \exp (V_j') \ , \qquad (11)$$

où $\exp (V_j')$: introduction d'une nouvelle alternative.

Dans ces cas, le dogit demeure distinct du logit, mais l'introduction d'une nouvelle alternative ou la variation des attributs de l'une d'elle, ne changera en rien la probabilité relative de choix des alternatives considérées.

C. Possibilité d'un "effet de revenu"

Les coefficients θ_i de (1) peuvent être interprétés comme indices de pondération des choix de la $i^{ième}$ alternative selon l'influence des attributs de toutes les alternatives. Selon Gaudry et Dagenais (avril 1978) les θ réagissent comme un "effet de revenu" qui peuvent accroître ou modérer "l'effet de substitution" automatique que l'on retrouve dans le modèle logit, par les cas suivants :

i - les θ_i réduisent η_{ii} et augmentent η_{ij} par rapport à ce que le format logit aurait donné;

ii - cette différence augmente avec l'importance des alternatives pour lesquelles les caractéristiques sont modifiées,

où

$$\eta_{ii} = B_k \, X_{ik} \, (X_{ik} + \mu_k)^{\lambda_k - 1} \left[\frac{(1 + \theta_i) \exp (V_i)}{\exp (V_i) + \theta_i \sum_j \exp (V_j)} - \frac{\exp (V_i)}{\sum_j \exp (V_j)} \right] \qquad (12)$$

= élasticité propre

et

$$\eta_{ij} = B_k X_{jk} (X_{jk} + \mu_k)^{\lambda_k - 1} \left[\frac{\theta_i \exp(V_j)}{\exp(V_i) + \theta_i \sum_j \exp(V_j)} - \frac{\exp(V_j)}{\sum_j \exp(V_j)} \right] \quad (13)$$

= élasticité croisée.

D. Intégration dans une même distribution d'un comportement contraint et libre

Réécrivons l'équations (1) :

$$p_i = \frac{\theta_i}{1 + \sum_j \theta_j} + \frac{\exp(V_i)}{(1 + \sum_j \theta_j) \sum_j \exp(V_j)} \quad (14)$$

On peut penser que, si le consommateur, selon l'approche de Stone (1954) dans la théorie de la demande, accorde une partie de son revenu à un $i^{ième}$ bien $(\theta_i / 1 + \sum_j \theta_j)$ pour satisfaire ses besoins essentiels sans tenir compte du prix du bien, alors $1 - \sum_i (\theta_i / 1 + \sum_j \theta_j)$ ou $1/(1 + \sum_j \theta_j)$ est la part du revenu libre qui reste. Si nous supposons que le logit explique correctement l'affectation du revenu "libre", la dépense pour la part du $i^{ième}$ bien va être la somme des éléments contraints et libres ou bien

$$p_i = \frac{\theta_i}{1 + \sum_j \theta_j} + \frac{1}{1 + \sum_j \theta_j} \cdot \frac{\exp(V_i)}{\sum_j \exp(V_j)} \quad (15)$$

qui est identique à (14).

Cette possibilité du dogit de réunir le comportement libre et contraint, peut-être interprétée comme étant un certain montant de "captivité" aux alternatives. Cette interprétation du dogit a été récemment développée par Ben-Akiva (1977). McFadden (1978) a récemment interprété le dogit comme un cas particulier de modèle de probabilités mixtes à utilité aléatoire ("Mixed probability random utility models").

4. Application du modèle à une série chronologique

Nous ferons des tests du modèle à partir de données agrégées sur les modes de paiement des déplacements par transport en commun sur le territoire de la Commission de Transport de la Communauté Urbaine de Montréal. Les usagers ont le choix d'acheter des billets relativement chers en petites quantités et des billets moins chers en grandes quantités. La probabilité d'utiliser un mode de paiement plutôt qu'un autre ne dépend pas seulement de la valeur unitaire et de l'investissement nécessaire mais aussi du motif de transport ou des activités économiques pour lesquelles on se déplace. Ces activités sont de nature cyclique et nous les représenterons par 12 "dummy" mensuelles. Nous associerons arbitrairement ces variables mensuelles avec les billets chers*. L'échantillon est une suite de 80 observations pour la période de mai 1965 à décembre 1971. Nous travaillons avec des données en dollars constants. Comme coefficient de dégonflement de l'inflation, nous utilisons l'indice des prix à la consommation, de Statistique Canada, pour la période considérée.

---

*Dans un choix binaire, on aurait pu sans affecter les résultats, les associer à l'autre alternative.

5. Méthode d'estimation

A. Nouvel arrangement pour les besoins de l'estimation

En présence de deux alternatives, le logarithme du rapport de la part de p_1 et de $p_2 = 1 - p_1$ donne la forme suivante en combinant (1) et (4) :

$$\ln \left[\frac{p_1 - (\theta_1/1 + \theta_1 + \theta_2)}{1 - p_1 - (\theta_2/1 + \theta_1 + \theta_2)} \right] = V_t + u_t \qquad (16)$$

où

$$V_t = \sum_k B_k \left((X_{1kt} + \mu_k)^{(\lambda_k)} - (X_{2kt} + \mu_k)^{(\lambda_k)} \right) \qquad (17)$$

t étant une observation 1, ..., N et u_t le terme d'erreur.

Dans les pages qui suivent, nous examinons les modèles dogit et logit avec des contraintes diverses sur les valeurs possibles des θ, λ et μ en (16) et (17).

Nous avons vu en (1) que $\theta_i \geq 0$. Ici en (16) on voit que pour ne pas avoir un logarithme négatif θ_1 et θ_2 doivent respecter les équations suivantes :

$$\theta_1/1 + \theta_1 + \theta_2 < \min (p_1) \qquad (18)$$

$$\theta_2/1 + \theta_1 + \theta_2 < \min (1 - p_1) \qquad (19)$$

Pour μ_k, nous avons vu en (6) que $X_{ik} + \mu_k > 0$. λ_k peut prendre des valeurs positives et négatives mais nous nous attendons à une valeur faible.

Pour simplifier l'estimation, vu le grand nombre de valeur que les paramètres peuvent prendre, les éléments du vecteur $\lambda = (\lambda_1, \ldots, \lambda_k)$ vont être contraints égaux.

Voici les principales formes estimées :

a) Dogit sans contrainte ou modèle Dsc.

θ_1 et θ_2 sont positifs et respectent les conditions de (18) et (19).

a') Dogit sans contrainte sauf que les éléments θ sont égaux ou modèle Dsce.

b) Dogit linéaire ou modèle Dℓ.

Comme en a sauf $\lambda = 1$ et $\mu = 0$.

c) Dogit multiplicatif ou modèle Dm.

Comme en a sauf $\lambda = \mu = 0$.

d) Logit sans contrainte ou modèle Lsc.

Comme en a sauf $\theta = 0$.

e) Logit linéaire ou modèle Ll.

Comme en d sauf $\lambda = 1$ et $\mu = 0$.

f) Logit multiplicatif ou modèle Lη.

Comme en d sauf $\lambda = \mu = 0$.

B. Fonction de vraisemblance en présence d'autocorrélation multiple

Si le vecteur du terme d'erreur de (16) est distribué normalement avec une moyenne zéro et une variance σ^2, la vraisemblance des observations de la part du marché p_1 peut être formulée pour des valeurs de μ et λ fixes comme étant :

$$\mathcal{L}(p_i \; ; \; B \; , \; \sigma^2 \; , \; \theta) = (2\pi\sigma^2)^{-N/2} \; . \; \exp\left(\frac{-u'u}{2\sigma^2}\right) |J \; (\theta \; ; \; p_1)| \qquad (20)$$

où $J|(\theta \; ; \; p_1)|$ est la valeur absolue du Jacobien de la transformation de la variable dépendante de (16) aux vraies valeurs observées de p_1, ou bien

$$|J \; (\theta \; ; \; p_1)| = \prod_{t=1}^{N} \left\{ (1 + \theta_1 + \theta_2) \left[p_{1t} - \frac{\theta_1}{1 + \theta_1 + \theta_2} \right] \left[1 - p_{1t} - \frac{\theta_2}{1 + \theta_1 + \theta_2} \right] \right\}^{-1}$$

$$= \prod_{t=1}^{N} J_t \qquad (21)$$

Il est pratique de maximiser la fonction du log concentré de la vraisemblance qui est

$$L \; (p_1 \; , \; \theta) = \frac{-N}{2} \left[\ell n \; (2\pi) + 1 \right] - \frac{N}{2} \ell n \left[\frac{u'u}{N} \right] + \sum_{t=1}^{N} \ell n \; J_t \qquad (22)$$

où

$$u_t = \ell n \left[\frac{p_1 - (\theta_1 / 1 + \theta_1 + \theta_2)}{1 - p_1 - (\theta_2 / 1 + \theta_1 + \theta_2)} \right] - V_t \; , \qquad (23)$$

que nous pouvons écrire

$$u_t = D(\theta)_t - \hat{V}(\lambda, \mu)_t \qquad (24)$$

Si le terme d'erreur de (16) suit un schème autorégressif du premier ordre, alors :

$$u_t = \rho_1 u_{t-1} + e_t \qquad (25)$$

et le log de la vraisemblance concentré va devenir

$$L(p_1 ; \theta, \rho_1) = -\frac{N}{2}\left[\ell n (2\pi) + 1\right] - \frac{N}{2} \ell n \left[\frac{e'e}{N}\right] + \frac{1}{2} \ell n (1 - \rho_1^2) + \sum_{t=1}^{N} \ell n J_t \qquad (26)$$

où e'e = u'Q'Qu et la matrice Q a la définition suivante :

$$Q = \begin{bmatrix} \sqrt{1-\rho_1^2} & 0 & \cdots & 0 \\ -\rho_1 & 1 & 0 \cdots & 0 \\ 0 & -\rho_1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots 0 & -\rho_1 \quad 1 \end{bmatrix} \qquad (27)$$

Si nous supprimons la première observation de l'échantillon, (26) devient :

$$L(p_1 ; \theta_1 \rho_1) = -\frac{N-1}{N}\left[\ell n (2\pi) + 1\right] - \frac{N-1}{2} \ell n \left[\frac{e'e}{N-1}\right] + \sum_{t=2}^{N} \ell n J_t \qquad (28)$$

En présence d'un ordre d'autocorrélation supérieur,

$$\mu_t = \sum_{\ell=1}^{R} \rho_\ell \, u_{t-\ell} + e_t \, , \qquad \rho = (\rho_1, \ldots, \rho_\ell, \ldots, \rho_R) \, , \tag{29}$$

et l'extention de (26) est :

$$L\,(p_1\,;\theta,\,\rho) = -\frac{N-R}{2}\left[\ell n\,(2\pi) + 1\right] - \frac{N-R}{2}\,\ell n\left[\frac{e'e}{N-R}\right] + \sum_{t=1+R}^{N} \ell n\, J_t \tag{30}$$

Cette vraisemblance simplifiée contourne les difficultés de programmation associées à l'incorporation du schème autorégressif généralisé à un coût d'une certaine perte d'efficacité Beach et MacKinnon (1978) par la réduction du nombre de données de l'ordre le plus élevé d'autocorrélation. Une discussion plus détaillée est fournie dans Gaudry et Wills (1977).

C. Technique de maximisation

La fonction de vraisemblance (30) est maximisée par une méthode non linéaire ou plus simplement par un balayage des valeurs possibles. Nous avons identifié le schème autorégressif de (29) comme étant du premier, du huitième et du douzième ordre en utilisant une technique qui est en accord avec l'esprit de Box-Jenskins. Nous avons appliqué les moindres carrés à (23) pour $\theta = \mu = 0$ et $\lambda = 1$. Cette spécification (pour λ, θ et μ fixe) a déterminé une équation non linéaire obtenue en combinant (24) et (29) :

$$D\,(\theta)_t - \sum_{\ell=1}^{R} D\,(\theta)_{t-\ell} = \hat{V}\,(\lambda,\,\mu)_t - \sum_{\ell=1}^{R} \hat{V}\,(\lambda,\,\mu)_{t-\ell} + e_t \, , \tag{31}$$

$$t = 1 + R, \ldots, N \, ,$$

que l'on a résolu par la méthode itérative de Cochrane-Orcutt (1949). Les valeurs de la fonction de vraisemblance (30) sont obtenues par le balayage pour les valeurs admissibles de θ, λ et μ sans changer le schème autorégressif.

6. Résultats

A. Mesure d'évaluation des résultats

Comme mesure de comparaison pour les différents modèles selon les valeurs de θ, λ et μ, nous allons utiliser la valeur du maximum de vraisemblance. Parmi le grand nombre de test conditionnel et inconditionnel sur θ, λ, B, μ et ρ, nous limiterons nos comparaisons, aux tests statistiques suivants :

T.1; Dsc. Vs Lsc.

$$\hat{L}\,(\hat{\theta}, \hat{\lambda}, \hat{\mu}, \hat{B}, (\hat{\rho})) - \hat{L}\,(\theta = 0, \hat{\lambda}, \hat{\mu}, \hat{B}, (\hat{\rho}))\ , \qquad (32)$$

T.2; Dm. Vs Lm.

$$\hat{L}\,(\hat{\theta}, \lambda = 0, \mu = 0, \hat{B}, (\hat{\rho})) - \hat{L}\,(\theta = 0, \lambda = 0, \mu = 0, \hat{B}, (\hat{\rho}))\ , \qquad (33)$$

T.3; Dl. Vs Ll;

$$\hat{L}\,(\hat{\theta}, \lambda = 1; \mu = 0, \hat{B}, (\hat{\rho})) - \hat{L}\,(\theta = 0, \lambda = 1, \mu = 0, \hat{B}, (\hat{\rho}))\ , \qquad (34)$$

T.4; Dsce. Vs Lsc.

$$\hat{L}\,(\hat{\theta}_1 = \hat{\theta}_2, \hat{\lambda}, \hat{\mu}, \hat{B}, (\hat{\rho})) - \hat{L}\,(\theta = 0, \hat{\lambda}, \hat{\mu}, \hat{B}, (\hat{\rho}))\ , \qquad (35)$$

T.1. compare le dogit au logit sans contrainte sur les autres coefficients alors que T.2 et T.3 comparent le dogit au logit mais avec une contrainte sur la forme fonctionnelle. T.4 compare le dogit au logit sans

contrainte sur les autres coefficients mais en imposant une condition supplémentaire au dogit; les θ doivent être égaux. Le nombre de degré de liberté associé à chacun des tests est égal à 2 sauf pour le T.4 dont le degré de liberté est égal à 1.

Les trois premiers tests vont être :

$$L\left[\Omega\right] - L\left[\omega\right] > \frac{1}{2}\chi^2_{\alpha} \quad (2) \qquad (36)$$

où Ω est le coefficient d'espace sous l'hypothèse constante, ω est l'espace restreint sous une hypothèse nulle et χ^2_{α} (2) est la distribution du Khi-carré pour 2 degrés de liberté avec un niveau de signification α. Le pourcentage des points de cette distribution avec 2 degrés de liberté est :

Niveau	$\chi^2_{.995}$	$\chi^2_{.99}$	$\chi^2_{.975}$	$\chi^2_{.95}$	$\chi^2_{.90}$	$\chi^2_{.75}$	$\chi^2_{.50}$
Valeur	10.6	9.21	7.38	5.99	4.61	2.77	1.39

Pour 3 degrés de liberté :

Valeur	12.8	11.3	9.35	7.81	6.25	4.11	2.37

B. Analyse des résultats

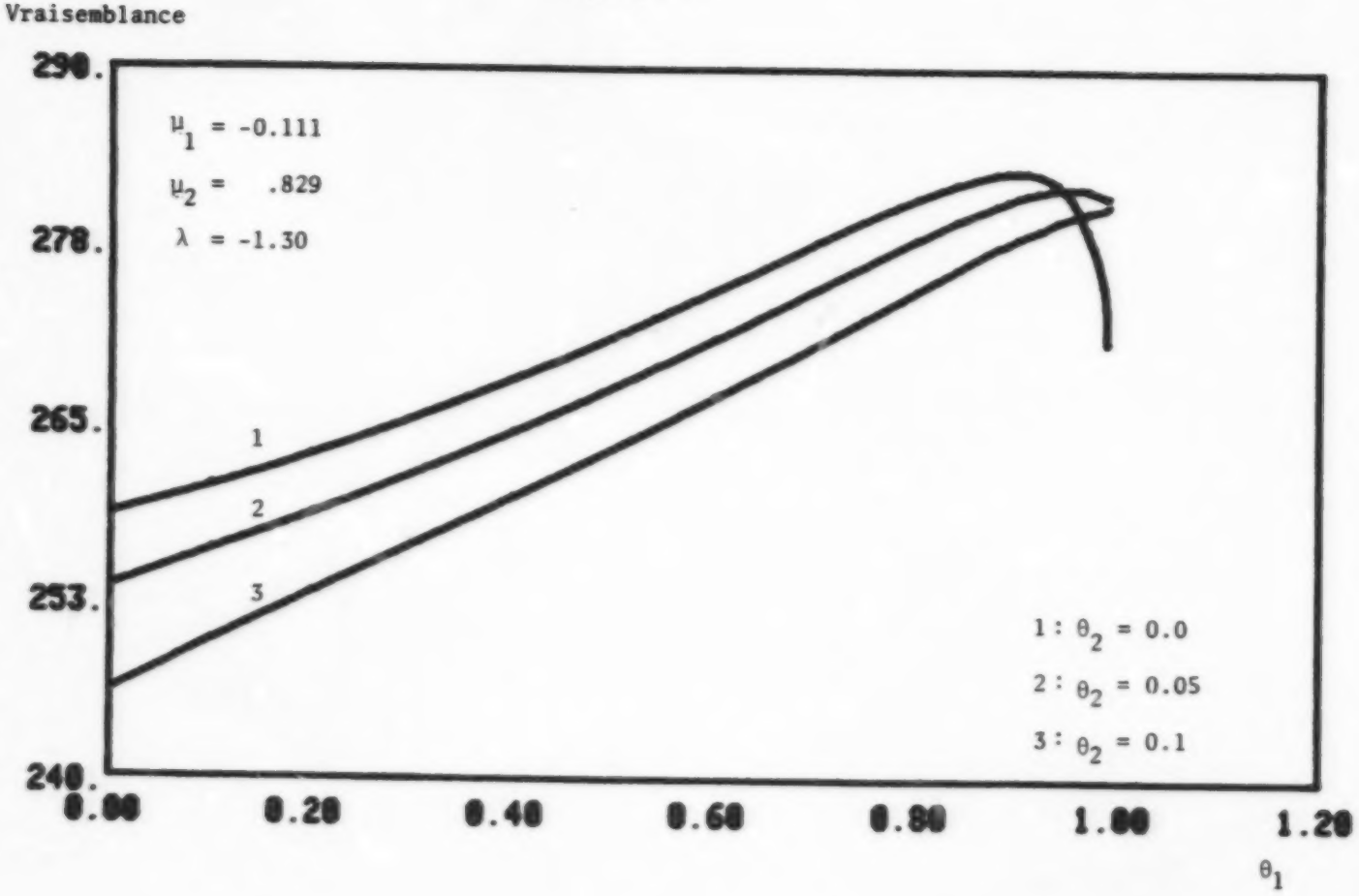
B.1. Influence générale des paramètres du dogit

A l'aide des figures qui suivent, nous pouvons clairement remarquer l'influence des paramètres $\theta = (\theta_1, \theta_2)$ du dogit sur les valeurs du log de la vraisemblance. Dans le graphique 1, nous observons le dogit sans contrainte. Nous avons fixé μ_1, μ_2 et λ pour celles qui maximisent la valeur de la vraisemblance du Dsc. La courbe A représente la valeur de θ_2 qui nous donne la plus grande vraisemblance. Au point $\theta_1 = 0$ nous obtenons le logit. Les courbes 2 et 3 donnent aussi le log de vraisemblance mais pour des valeurs de $\theta_2 \neq 0$.

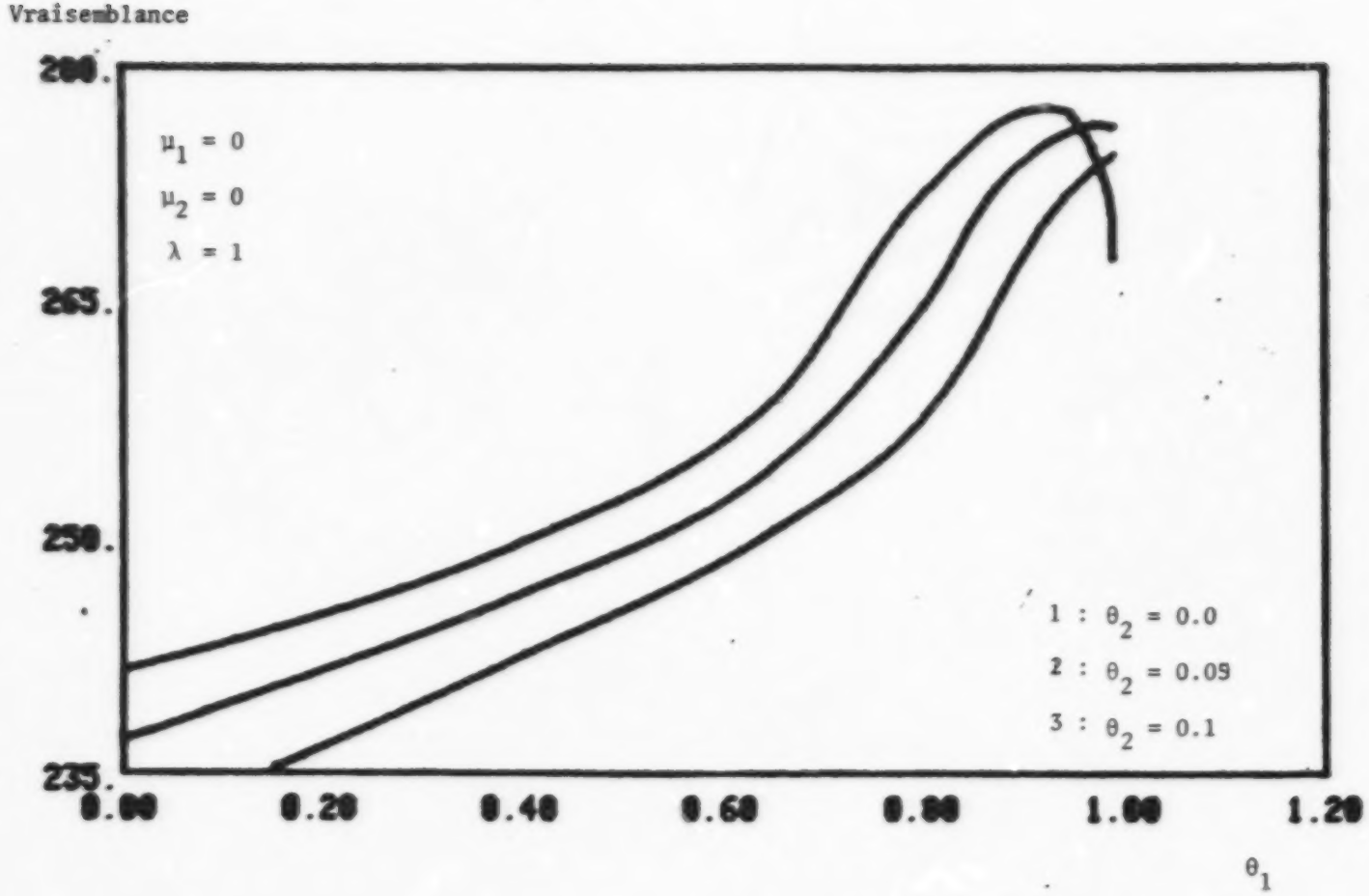
Nous devons faire une réserve à l'analyse de ces courbes, car une variation du point optimum, dans notre cas $\theta_1 = 0.901$ et $\theta_2 = 0$, entraîne un déplacement de la valeur μ_1, μ_2 et λ associé à la nouvelle paire de θ et le log de la vraisemblance peut possiblement être améliorée pour d'autres valeurs de μ et λ. Dans les graphiques 2 et 3, nous n'avons pas ce phénomène car les μ sont nuls par hypothèse et λ représente la forme linéaire si = 1 ou logarithmique si = 0 comme nous avons vu précédemment.

Une particularité intéressante des graphiques est la croissance unique jusqu'à l'optimum du log de la vraisemblance, puis la décroissance de celle-ci. Il semble apparent par les graphiques 1, 2 et 3 que l'introduction des θ va améliorer le log de la vraisemblance et aussi les résultats du logit.

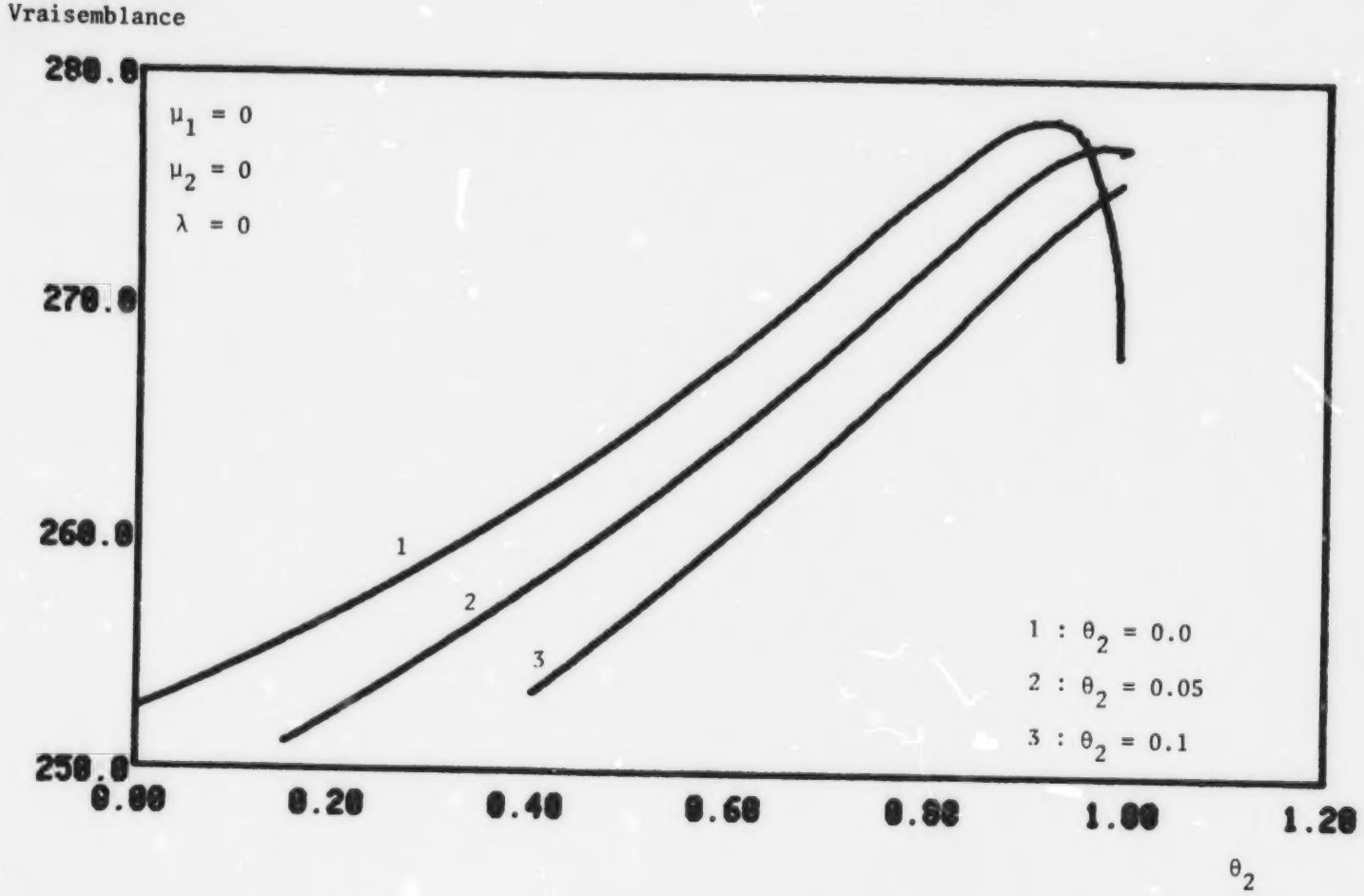
Graphique 1


Vraisemblance
290.
278.
265.
253.
240.
μ_1 = -0.111
μ_2 = .829
λ = -1.30
1
2
3
1 : θ_2 = 0.0
2 : θ_2 = 0.05
3 : θ_2 = 0.1
0.00
0.20
0.40
0.60
0.80
1.00
1.20
θ_1

Graphique 2


Vraisemblance
270.
265.
260.
255.
$\mu_1 = 0$
$\mu_2 = 0$
$\lambda = 1$
1 : $\theta_2 = 0.0$
2 : $\theta_2 = 0.05$
3 : $\theta_2 = 0.1$
0.00
0.20
0.40
0.60
0.80
1.00
1.20
θ_1

Graphique 3


Vraisemblance
280.0
270.0
260.0
250.0
$\mu_1 = 0$
$\mu_2 = 0$
$\lambda = 0$
1
2
3
1 : $\theta_2 = 0.0$
2 : $\theta_2 = 0.05$
3 : $\theta_2 = 0.1$
0.00
0.20
0.40
0.60
0.80
1.00
1.20
θ_2

B.2. Comparaison entre les différents modèles

Les résultats qui nous intéressent pour les différents modèles étudiés sont au tableau 1. L'analyse de ceux-ci nous révèlent :

i) la valeur du log de vraisemblance est nettement supérieure pour les modèles dogit comparés aux logit de même force. Les tests (T.1; Dsc Vs Lsc, T.2; D1 Vs L1, T.3; Dm Vs Lm) donnent respectivement 22.160, 29.782 et 30.687. Nous pouvons donc rejeter avec un haut niveau de confiance le logit et accepter le dogit;

ii) l'effet de captivité, comme décrit précédemment en 3.D. associé aux billets chers qui demandent un petit investissement est fort élevé; environ 47%. Ceci peut signifier que la liquidité nécessaire et le risque de perte des billets achetés en grande quantité peuvent être des contraintes pour l'usager moyen du transport en commun;

iii) dans le cas du dogit où nous contraignons les θ égaux entre eux, nous obtenons $\theta_1 = \theta_2 = 0$ ou le logit sans contrainte pour la grande valeur du log de la vraisemblance comme nous pouvons voir dans la colonne Dc/Lsc. Nous obtenons une différence nulle pour le T.4 d'où l'importance de ne pas mettre de restrictions sur les θ. Ce résultat n'est pas surprenant pour une situation de deux altervatives; une valeur positive pour θ_1 et θ_2 égale aurait impliqué, selon l'équation 10, une attraction égale de chacune des alternatives et une part constante du marché égale. Ce résultat n'est pas trivial pour une paire d'alternatives dans une situation de plusieurs choix;

iv) si nous divisons les modèles en deux classes, dogit et logit, le dogit et le logit sans contrainte domine les modèles où nous imposons une forme fonctionnelle a priori. La valeur du log de la vraisemblance est significative dans le dogit à un niveau un peu plus élevé que 75% alors que pour le logit ce niveau est un peu plus élevé que 90%. Ce résultat diffère de l'analyse de Gaudry et Wills (1978) où ceux-ci travaillent en dollars courants;

v) en ce qui concerne la valeur des résultats pour le signe des coefficients, leur student et leur élasticité; ceux-ci sont plutôt pauvres. A priori, nous pouvions croire que les coefficients du tarif et de l'investissement seraient négatifs avec des élasticités faibles. Dans tous les modèles, le signe d'une variable est contraire aux attentes et les élasticités associées à ceux-ci varient beaucoup;

vi) un fait intéressant que nous avons noté; nous avons posé le schème autorégressif 1, 8 et 12 par les résultats préliminaires sur le logit linéaire et multiplicatif. A ce moment, on ne voyait pas le bien-fondé du schème 8 bien que celui-ci améliorait nos résultats. Hors, le dogit sans contrainte semble impliquer un schème autorégressif 1 et 12, le student de 8 étant très faible. Nous n'avons pas cru bon de refaire l'estimation avec le schème 1, 12 vu le coût et le temps nécessaire pour ce, avec une faible possibilité d'améliorer les résultats sensiblement.

Tableau 1

Résultats de la série chronologique agrégée en dollars constants

Paramètres	Type de modèle*					
	Dsc	D1	Dm	Dc/Lsc	L1	Lm
0_1 billets ch.	0.901	0.917	0.922	0.0	0.0	0.0
0_2 billets - ch.	0.0	0.0	0.0	0.0	0.0	0.0
μ_1 prix	-0.111	0.0	0.0	-0.132	0.0	0.0
μ_2 investissement	0.829	0.0	0.0	0.861	0.0	0.0
λ prix	-1.30	1.0	0.0	-0.94	1.0	0.0
investissement	-1.30	1.0	0.0	-0.94	1.0	0.0
Coefficient prix	-0.991	67.241	22.429	-0.629	28.429	9.799
t-statistique	(-6.185)	(1.650)	(3.332)	(-8.353)	(1.792)	(3.796)
élasticité**	-6.36	0.863	0.751	-4.871	2.938	4.863
	7.14	-0.836	-0.751	5.505	-2.848	-4.863
Coefficient inv.	58.658	-0.544	-2.918	23.043	-0.238	-1.276
t-statistique	(9.742)	(-1.963)	(-2.962)	(11.884)	(-2.086)	(-3.131)
élasticité**	2.75	-0.026	-0.098	2.210	-0.092	-0.633
	-1.12	0.128	0.098	-1.363	0.451	0.633
Valeur ρ_1	0.603	0.756	0.778	0.119	0.525	0.565
t-statistique	(6.217)	(9.394)	(9.899)	(1.037)	(5.423)	(5.990)
Valeur ρ_8	0.074	0.240	0.229	0.171	0.380	0.370
t-statistique	(0.768)	(2.637)	(2.504)	(1.796)	(3.878)	(3.758)
Valeur ρ_{12}	-0.221	-0.050	-0.051	-0.300	-0.039	-0.053
t-statistique	(-2.312)	(-0.059)	(-0.619)	(-3.478)	(-0.482)	(-0.656)
$L(\Omega)$	282.703	278.578	278.012	260.543	252.921	252.016
$L(\Omega) - L(\omega)$	0.0	4.125	4.691	22.160	20.782	30.687
R^2 sur $D(\rho)$***	0.816	0.507	0.496	0.885	0.510	0.496
R^2 sur D***	0.930	0.921	0.920	0.886	0.857	0.853
D.L.	0	3	3	1/2	5	5

* Sens des lettres : D = dogit, L = logit, sc = sans contrainte, 1 = linéaire, m = multiplicatif. Le t-statistique est calculé à chaque modèle pour θ, λ et μ donné.

** L'élasticité propre (numérateur) et croisée (dénominateur) de l'alternative 1 (billets chers) est donnée.

*** Calculé sur les variables transformées par λ et μ.

UNIVERSITÉ DE MONTRÉAL

7. Conclusion

Nous avons voulu montrer par ce rapport de recherche, l'avantage du dogit sur le logit. Nous avons obtenu des résultats qui nous permettent de rejeter le logit avec un haut niveau de confiance sans tenir compte de la forme fonctionnelle a priori ou estimée. L'utilisation de dollars constants, n'a pas donné de résultats supérieurs à ceux de Gaudry et Wills (1978) qui ont utilisé des dollars courants.

Ce travail n'est qu'une première recherche d'estimation du dogit. Par l'amélioration de la spécification du modèle, par le retrait de la contrainte où les λ sont fixes et par une meilleure détermination du schème autorégressif, on pourrait peut-être améliorer nos résultats. Nous ne croyons pas toutefois que nos résultats seraient profondément modifiés.

UNIVERSITÉ DE MONTRÉAL

Remerciements

Je tiens à offrir mes recerciements à Monsieur Marc Gaudry, pour ses nombreux conseils et de sa disponibilité pour la direction de ce rapport de recherche. Monsieur Léonard Dudley, pour avoir accepter d'être second correcteur, Danielle Bilodeau pour avoir accepté de programmer ce type de modèle, Monsieur Marcel Dagenais, pour les conseils qu'il m'a donnés, le Centre de calcul de l'Université de Montréal, pour son assistance technique ainsi que tous mes amis, pour leurs encouragements.

Bibliographie

BEACH, C.M. et MACKINNON, J.G., "A maximum likelihood procedure for regression with autocorrelated errors", Econometrica, vol. 46, 1978, p. 51-58.

BEN-AKIVA, M., "Choice models with simple choice set generating processes" Department of Civil Engineering, M.I.T., 1977.

BOX, G.E.P. et COX, D.R., "An analysis of transformations", Journal of the Royal Statistical Society, série B, 1964, p. 211-243.

BOX, G.E.P. et JENKINS, G.M., Time-Series Analysis, Forecasting and Control, Holden-Day, San Francisco, 1970.

COCHRANE, D. et ORCUTT, G.H., "Application of least-squares regression to relationships containing autocorrelated error terms", Journal of the American Statistical Association, vol. 44, 1949, p. 32-61.

DOMENCICH, Thomas A. et McFADDEN, Daniel, Urban Travel Demand, North-Holland, Amsterdam, 1975.

GAUDRY, Marc J.I. et DAGENAIS, Marcel G., "The dogit model", Centre de Recherche sur les Transports, Université de Montréal, publication -82, avril 1978.

GAUDRY, Marc J.I. et WILLS, Michael J., "Estimating the functional form of travel demand models", Centre de Recherche sur les Transports, Université de Montréal, publication -63, juin 1977.

GAUDRY, Marc J.I. et WILLS, Michael J., "Testing the dogit model with aggregate time-series and cross-setional travel data", Centre de Recherche sur les Transports, publication #94 (préliminaire), mars 1978.

LANCASTER, Kevin J., "A new approach to consumer theory", Journal of Political Economy, vol. 64, 1966, p. 132-157.

LUCE, R.D., Individual Choice Behavior, Wiley, New York, 1959.

McFADDEN, Daniel, Frontiers in Econometrics, édition P. Zarembka, Academic Press, New York, chapitre 4, 1973, p. 105-123.

SCHLESSELMAN, J., "Power families : a note on the Box and Cox transformation", Journal of the Royal Statistical Society, série B, vol. 33, 1971.

STONE, R., "Linear expenditures systems and demand analysis : an application to the pattern of Britsh demand", The Economic Journal, vol. 64, 1954, p. 511-527.

TUKEY, J.W., "On the comparative anatomy of transportation", Annals of Mathematical Statistics, vol. 28, 1957, p. 602-632.